

5

UNITED STATES PATENT APPLICATION

10

of

Brendan LARDER, Sharon KEMP, Stuart BLOOR and Ann BROPHY

15

FOR

**METHOD FOR MUTATION DETECTION IN HIV-1
USING POL SEQUENCING**

**METHOD FOR MUTATION DETECTION IN HIV-1 USING POL
SEQUENCING**

FIELD OF THE INVENTION

Ins. C1 C1
The present invention relates to a method for detecting mutations within the HIV
5 *pol* gene of HIV-1 isolates and in particular with the design of amplification primers and
sequencing primers for use in the analysis of the coding domains for the protease and
reverse transcriptase, respectively.

BACKGROUND OF THE INVENTION

10 The availability of rapid, high-throughput automated DNA sequencing technology
has obvious applications in clinical research, including the detection of variations in virus
populations and mutations responsible for drug resistance in virus genomes. However,
analysis of clinical samples by manual sequencing or polymerase chain reaction-(PCR)
based point mutation assays has revealed that complex mixtures of wild type and mutant
HIV-1 genomes can occur during drug therapy. Therefore, to assess the likely
15 susceptibility of a virus population to a particular drug therapy, it would be desirable to
perform DNA sequence analysis that can simultaneously quantitate several resistance
mutations in multiple genomes. A particular advantage of analyzing the sequence of
more than one *pol* gene enzyme (Protease and Reverse transcriptase) is that the studied
material reflects to a greater extent the viral genetic diversity in the particular patient
20 being investigated.

SUMMARY OF THE INVENTION

The aim of the present invention is thus to provide a reliable sequence analysis
method and kit for performing mutation analysis of the *pol* gene of HIV-1 virus isolates.

In one embodiment, the present invention relates to a method for mutation
25 analysis of the HIV *pol* gene of a HIV-1 virion comprising the steps of:
a) isolation of a sample comprising HIV-1 RNA,
b) amplifying RNA using outer primers as represented in SEQ ID No: 1
(OUT3) and 2 (PRTO-5),

- c) amplifying the product of (b) using a 5' and 3' primer chosen from the inner primers as represented in SEQ ID No: 3 (PCR2.5), 4 (PCR2.3), 5 (SK107) and 6 (SK108), and
- d) sequencing this secondary obtained product.

5 The present invention also provides a method for mutation analysis of the HIV *pol* gene of HIV-1 isolates comprising the steps of:

- a) isolation of a sample comprising HIV-1 DNA,
- b) amplifying DNA using outer primers as represented in SEQ ID No: 1 (OUT3) and 2 (PRTO-5),
- 10 c) amplifying the product of (b) using a 5' and 3' primer chosen from the inner primers as represented in SEQ ID No: 3 (PCR2.5), 4 (PCR2.3), 5 (SK107) and 6 (SK108), and
- d) sequencing this secondary obtained product.

 The present invention also relates to a primer as described herein (see Table 1) and used to analyze the sequence of the HIV *pol* gene of HIV-1 isolates. In a further
15 embodiment, the present invention relates to a diagnostic kit for the mutation analysis of the HIV *pol* gene of HIV-1 isolates comprising at least one of the primers as shown in Table 1.

 Additional objects and advantages of the invention will be set forth in part in the
20 description which follows, and in part will be apparent from the description, or may be learned by practice of the invention. The objects and advantages of the invention will be realized and attained by means of the elements and combinations particularly pointed out in the appended claims.

 It is to be understood that both the foregoing general description and the following
25 detailed description are exemplary and explanatory only and are not restrictive of the invention as claimed.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1: Schematic overview of the total coding region of the protease- RT coding domain of HIV-1 isolates. The length in nucleotides of both coding regions is indicated. Regions that are sequenced using respectively mentioned sequencing primers are shown. Primary sequences and the secondary sequences are schematically presented.

DETAILED DESCRIPTION OF THE INVENTION

The present invention, in one aspect, relates to a method for mutation analysis of the HIV *pol* gene of a HIV-1 virion comprising the steps of:

- a) isolation of a sample comprising HIV-1 RNA,
- 10 b) PCR amplifying RNA using outer primers as represented in SEQ ID No: 1 (OUT3) and 2 (PRTO-5),
- c) PCR amplifying said PCR product using a 5' and 3' primer chosen from the inner primers as represented in SEQ ID No: 3 (PCR2.5), 4 (PCR2.3), 5 (SK107) and 6 (SK108), and
- 15 d) sequencing this secondary obtained PCR product.

In a preferred embodiment, the amplifying is via nested PCR. The secondary obtained PCR product may be sequenced using at least one sequencing primer chosen from any of SEQ ID No: 7 to 12 (Seq1FOR, Seq2FOR, Seq3F, Seq1B, Seq3B, Seq6R, Seq1F, Seq2A, Seq3A, Seq5A, Seq7A, Seq2B, Seq4B, Seq6B, Seq7B, Seq4A, Seq6A, Seq5B; see Table 1). In one embodiment, RNA is viron RNA extracted from the sample. In another embodiment, RNA is cell RNA extracted from an infected cell sample.

The present invention describes a mutation analysis of the *pol* gene of HIV-1 isolates including group M and group O viruses, in particular group M viruses. Mixed populations carrying mutations can be detected when present down to at least 20%.

25 The present invention also provides a method for mutation analysis of the HIV *pol* gene of HIV-1 isolates comprising the steps of:

- a) isolation of a sample comprising HIV-1 DNA,
- b) PCR amplifying DNA using outer primers as represented in SEQ ID No: 1

(OUT3) and 2 (PRTO-5),

c) PCR amplifying said PCR product using a 5' and 3' primer chosen from the inner primers as represented in SEQ ID No: 3 (PCR2.5), 4 (PCR2.3), 5 (SK107) and 6 (SK108), and

5 d) sequencing this secondary obtained PCR.

In one embodiment, the amplifying is via nested PCR. The secondary obtained PCR product may be sequenced using at least one sequencing primer chosen from any of SEQ ID No: 7 to 12 (Seq1FOR, Seq2FOR, Seq3F, Seq1B, Seq3B, Seq6R, Seq1F, Seq2A, Seq3A, Seq5A, Seq7A, Seq2B, Seq4B, Seq6B, Seq7B, Seq4A, Seq6A, Seq5B; see Table

10 1). In one embodiment, DNA is viral DNA extracted from the isolated sample material.

According to a preferred method said secondary PCR product is sequenced using a primer as represented in SEQ ID No: 7 (Seq1FOR).

According to a preferred method said secondary PCR product is sequenced using a primer as represented in SEQ ID No: 8 (Seq2FOR).

15 According to a preferred method said secondary PCR product is sequenced using a primer as represented in SEQ ID No: 9 (Seq3F).

According to a preferred method said secondary PCR product is sequenced using a primer as represented in SEQ ID No: 10 (Seq 1B).

20 According to a preferred method said secondary PCR product is sequenced using a primer as represented in SEQ ID No:11 (Seq3B).

According to a preferred method said secondary PCR product is sequenced using a primer as represented in SEQ ID No: 12 (Seq6R).

25 The present invention also provides a method according to the present invention wherein one of the initial sequencing primers is replaced by one or a pair of replacement primers (Table 2). For example, if Seq2FOR (SEQ ID No: 8) failed it is replaced by Seq3A (SEQ ID No: 15) and Seq5A (SEQ ID No: 16). However in principle any described primer that obtains sequence from the region that Seq2FOR (SEQ ID No: 8) was expected to cover can be used i.e. Seq3A (SEQ ID No: 15), Seq4A (SEQ ID No: 22)

or Seq5A (SEQ ID No: 16) (see Figure 1). In addition, Seq6A (SEQ ID No: 23) and Seq5B (SEQ ID No: 24) were also not proposed to replace a specific initial primer but can be used to cover respective sequence domains (see Figure 1).

5 In preferred methods according to the present invention the initial sequencing primer as represented in SEQ ID No 7 (Seq 1FOR) is replaced by a primer set as represented in SEQ ID No: 13 (Seq1F) and 14 (Seq2A).

In preferred methods according to the present invention the initial sequencing primer as represented in SEQ ID No 8 (Seq2FOR) is replaced by a primer set as represented in SEQ ID No: 15 (Seq3A) and 16 (Seq5A).

10 In preferred methods according to the present invention the initial sequencing primer as represented in SEQ ID No 9 (Seq3F) is replaced by a primer set as represented in SEQ ID No: 16 (Seq5A) and 17 (Seq7A).

15 In preferred methods according to the present invention the initial sequencing primer as represented in SEQ ID No 10 (Seq 1B) is replaced by a primer set as represented in SEQ ID No: 4 (PCR2.3) and 18 (Seq2B).

In preferred methods according to the present invention the initial sequencing primer as represented in SEQ ID No 11 (Seq3B) is replaced by a primer set as represented in SEQ ID No: 18 (Seq2B) and 19 (Seq4B).

20 In preferred methods according to the present invention the initial sequencing primer as represented in SEQ ID No 12 (Seq6R) is replaced by a primer set as represented in SEQ ID No: 20 (Seq6B) and 21 (Seq7B).

Preferably, the methods according to present invention involve a sequencing step wherein said secondary PCR product is sequenced using a primer as represented in SEQ ID No 13 (Seq1F).

25 Preferably, the methods according to present invention involve a sequencing step wherein said secondary PCR product is sequenced using a primer as represented in SEQ ID No 14 (Seq2A).

Preferably, the methods according to present invention involve a sequencing step wherein said secondary PCR product is sequenced using a primer as represented in SEQ ID No 15 (Seq3A).

5 Preferably, the methods according to present invention involve a sequencing step wherein said secondary PCR product is sequenced using a primer as represented in SEQ ID No 16 (Seq5A).

Preferably, the methods according to present invention involve a sequencing step wherein said secondary PCR product is sequenced using a primer as represented in SEQ ID No 17 (Seq7A).

10 Preferably, the methods according to present invention involve a sequencing step wherein said secondary PCR product is sequenced using a primer as represented in SEQ ID No 18 (Seq2B).

15 Preferably, the methods according to present invention involve a sequencing step wherein said secondary PCR product is sequenced using a primer as represented in SEQ ID No 19 (Seq4B).

Preferably, the methods according to present invention involve a sequencing step wherein said secondary PCR product is sequenced using a primer as represented in SEQ ID No 20 (Seq6B).

20 Preferably, the methods according to present invention involve a sequencing step wherein said secondary PCR product is sequenced using a primer as represented in SEQ ID No 21 (Seq7B).

Preferably, the methods according to present invention involve a sequencing step wherein said secondary PCR product is sequenced using a primer as represented in SEQ ID No 22 (Seq4A).

25 Preferably, the methods according to present invention involve a sequencing step wherein said secondary PCR product is sequenced using a primer as represented in SEQ ID No 23 (Seq6A).

Preferably, the methods according to present invention involve a sequencing step wherein said secondary PCR product is sequenced using a primer as represented in SEQ ID No 24 (Seq5B).

A primer acts as a point of initiation for synthesis of a primer extension product that is complementary to the nucleic acid strand to be copied. The place of hybridization is determined by the primer- and target sequence. As known by the skilled person in the art, specificity of the annealing can be guaranteed by choosing a sequence domain within the target sequence, which is unique, compared to other non-target sequences.

Nevertheless, start and stop of the primer onto the target sequence may be located some nucleotides up- or downstream the defined primer site without interfering with this specificity.

Consequently, the present invention also provides a method as described above wherein the sequencing primer is chosen up to 1, 2, 3 or 4 nucleotides upstream or downstream the described primer region.

The present invention also provides a method as described above wherein the outer primer is chosen up to 1, 2, 3 or 4 nucleotides upstream or downstream the described primer region.

The present invention also provides a method as described above wherein the inner primer is chosen up to 1, 2, 3 or 4 nucleotides upstream or downstream the described primer region.

The present invention also provides a method as described above wherein the sample contains free virion particles or virus infected cells.

In particular, the present invention also provides a method as described above wherein the sample is any biological material taken either directly from the infected human being (or animal), or after culturing (enrichment). Biological material may be, e.g., expectorations of any kind, broncheolavages, blood (plasma, serum), skin tissue, biopsies, sperm, semen, lymphocyte blood culture material, colonies, liquid cultures, fecal samples, urine etc.

The present invention also relates to a primer as described above (see Table 1) and used to analyze the sequence of the HIV *pol* gene of HIV-1 isolates.

Preferentially, such methods according to the present invention involve the sequencing of the defined primary PCR product.

5 The present invention also relates to a diagnostic kit for the mutation analysis of the HIV *pol* gene of HIV-1 isolates comprising at least one of the primers as shown in Table 1. The following definitions serve to illustrate the terms and expressions used in the present invention.

10 The term "drug-induced mutation" means any mutation different from consensus wild-type sequence, more in particular it refers to a mutation in the HIV protease or RT coding region that, alone or in combination with other mutations, confers a reduced susceptibility of the isolate to the respective drug.

15 The term "target sequence" as referred to in the present invention describes the nucleotide sequence of the wild type, polymorphic or drug induced variant sequence of the protease and RT gene of HIV-1 isolates to be specifically detected by sequence analysis according to the present invention. This nucleotide sequence may encompass one or several nucleotide changes. Target sequences may refer to single nucleotide positions, nucleotides encoding amino acids or to sequence spanning any of the foregoing nucleotide positions. In the present invention said sequence often includes one or two
20 variable nucleotide positions.

It is to be understood that the complement of said target sequence is also a suitable target sequence in some cases.

25 The target material in the samples to be analyzed may either be DNA or RNA, e.g., genomic DNA, cDNA, messenger RNA, viral RNA or amplified versions thereof. These molecules are also termed polynucleic acids. It is possible to use DNA or RNA molecules from HIV samples in the methods according to the present invention.

Well-known extraction and purification procedures are available for the isolation of RNA or DNA from a sample, (e.g., in Maniatis et al., Molecular Cloning: A

Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory Press (1989), the disclosure of which is hereby incorporated by reference).

5 The term "primer" refers to single stranded sequence-specific oligonucleotide capable of acting as a point of initiation for synthesis of a primer extension product that is complementary to the nucleic acid strand to be copied. The length and the sequence of the primer must be such that they allow priming the synthesis of the extension products.

Preferentially, the primer is about 5-50 nucleotides long. Specific length and sequence will depend on the complexity of the required DNA or RNA targets, as well on the conditions of primer use such as temperature and ionic strength.

10 The fact that amplification primers do not have to match exactly with the corresponding template to warrant proper amplification is ample documented in the literature (Kwok, S., Kellog, D., McKinney, N., Spasic, D., Goda, L., Levenson, C. and Sinisky, J., Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency virus type 1 model studies, *Nucl. Acids Res.*, **18**, 999 (1990),
15 the disclosure of which is hereby incorporated by reference.).

The amplification method used can be either polymerase chain reaction (PCR; Saiki R, Walsh P, Levenson C, Erlich H., Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes, *Proc Natl Acad Sci USA*, **86**,6230-6234 (1989)), ligase chain reaction (LCR; Landgren, U; Kaiser, R; Sanders, J; Hood, L., A ligase-mediated gene detection technique, *Science*, **241**,1077-1080 (1988);
20 Wu, D; Wallace, B., The ligation amplification reaction (LAR)- amplification of specific DNA sequences using sequential rounds of template-dependent ligation. *Genomics*, **4**, 560-569 (1989); Barany, F., Genetic disease detection and DNA amplification using cloned thermostable ligase. *Proc. Natl. Acad Sci USA*, **88**,189-193 (1991)), nucleic acid
25 sequence-based amplification (NASBA; Guatelli, JC; Whitfield, KM; Kwoh, DY; Barringer, KJ, Richman, DD; Gingeras, TR., Isothermal, in vitro amplification of nucleic acids by a multienzyme reaction modeled after retroviral replication. *Proc. Natl. Acad. Sci USA*, **87**, 1874-1878 (1990); Compton, J., Nucleic acid sequence-based amplification.

Nature, **350**, 91-92 (1991)), transcription-based amplification system (TAS; Kwoh, D; Davis, G; Whitfield, K; Chappelle, H; Dimichele, L; Gingeras, T., Transcription-based amplification system and detection of amplified human immunodeficiency virus type 1 with a bead-based sandwich hybridization format, *Proc. Natl. Acad Sci USA*, **86**, 1173-1177 (1989)), strand displacement amplification (SDA; Duck, P., Probe amplifier system based on chimeric cycling oligonucleotides, *Biotechniques*, **9**, 142-147 (1990); Walker, G; Little, M; Nadeau, J; Shank, D., Isothermal in vitro amplification of DNA by a restriction enzyme/DNA polymerase system, *Proc. Natl. Acad Sci USA*, **89**, 392-396 (1992)) or amplification by means of Qss replicase (Lizardi, P; Guerra, C; Lomeli, H; Tussie-Luna, I; Kramer, F., Exponential amplification of recombinant RNA hybridization probes, *Bio/Technology*, **6**, 1197-1202 (1988); Lomeli, H; Tyagi, S; Printchard, C; Lisardi, P; Kramer, F., Quantitative assays based on the use of replicatable hybridization probes. *Clin. Chem.*, **35**, 1826-1831 (1989)) or any other suitable method to amplify nucleic acid molecules known in the art. The disclosures of the above listed references are hereby incorporated by reference.

The oligonucleotides used as primer may also comprise nucleotide analogues such as phosphothiates (Matsukura M, Shinozuka K, Zon G, Mitsuya H, Reitz M, Cohen J, Broder S., Phosphorothioate analogs of oligodeoxynucleotides: inhibitors of replication and cytopathic effects of human immunodeficiency virus, *Proc. Natl. Acad. Sci. USA*, **84**, 7706-10 (1987)), alkylphosphorothiates (Miller P, Yano J, Yano E, Carroll C, Jayaram K, Ts'o P., Nonionic nucleic acid analogues. Synthesis and characterization of dideoxyribonucleoside methylphosphonates, *Biochemistry*, **18**, 5134-43 (1979)) or peptide nucleic acids (Nielsen P, Egholm M, Berg R, Buchardt O., Sequence-selective of DNA by strand displacement with a thymine-substituted polyamide, *Science*, **254**, 1497-500 (1991); Nielsen P, Egholm M, Berg R; Buchardt O., Sequence specific inhibition of DNA restriction enzyme cleavage by PNA, *Nucleic-Acids-Res.*, **21**, 197-200 (1993)) or may contain intercalating agents (Asseline U, Delarue M, Lancelot G, Toulme F, Thuong N., Nucleic acidbinding molecules with high affinity and base sequence specificity:

intercalating agents covalently linked to oligodeoxynucleotides. *Proc. Natl. Acad. Sci. USA* **81**, 3297-301 (1984)). The disclosures of the above listed references are hereby incorporated by reference.

The figures, tables and examples as given below exemplify the present invention.

5 These data are not meant to limit the scope of the present invention.

a 7 ~~Table 1. Sequence of the amplification and sequencing primers used. Name and sequence identification numbers are indicated.~~

NAME	SEQUENCE	SEQ ID NO
cDNA synthesis and first round PCR		
OUT 3	5'-CAT-TGC-TCT-CCA-ATT-ACT-GTG-ATA-TTT-CTC-ATC-3'	SEQ ID 1
PRTO-5	5'-GCC-CCT-AGG-AAA-AAG-GGC-TGT-TGG-3'	SEQ ID 2
Second round (nested) PCR		
Set A		
PCR2.5	5'-CCT-AGG-AAA-AAG-GGC-TGT-TGG-AAA-TGT-GG-3'	SEQ ID 3
PCR2.3	5'-CTA-ACT-GGT-ACC-ATA-ATT-TCA-CTA-AGG-GAG-G-3'	SEQ ID 4
Set B		
SK107	5'-CAT-CTA-CAT-AGA-AAG-TTT-CTG-CTC-C-3'	SEQ ID 5
SK108	5'-CTA-GGA-AAA-AGG-GCT-GTT-GGA-AAT-G-3'	SEQ ID 6
Primary Sequencing primers		
Seq1FOR	5'-GAG-AGC-TTC-AGG-TTT-GGG-G-3'	SEQ ID 7
Seq2FOR	5'-AAT-TGG-GCC-TGA-AAA-TCC-3'	SEQ ID 8
Seq3F	5'-CCT-CCA-TTC-CTT-TGG-ATG-GG-3'	SEQ ID 9
Seq1B	5'-CTC-CCA-CTC-AGG-AAT-CC-3'	SEQ ID 10
Seq3B	5'-GTA-CTG-TCC-ATT-TAT-CAG-G-3'	SEQ ID 11
Seq6R	5'-CTT-CCC-AGA-AGT-CTT-GAG-TCC-3'	SEQ ID 12

a

Secondary sequencing primers		
Seq1F	5'-CAG-ACC-AGA-GCC-AAC-AGC-CCC-3'	SEQ ID 13
Seq2A	5'-CAC-TCT-TTG-GCA-ACG-ACC-C-3'	SEQ ID 14
Seq3A	5'-GGT-ACA-GTA-TTA-GTA-GGA-CC-3'	SEQ ID 15
Seq5A	5'-GTA-CTG-GAT-GTG-GGT-GAT-GC-3'	SEQ ID 16
Seq 7A	5'-GTG-GGA-AAA-TTG-AAT-TGG-G-3'	SEQ ID 17
PCR2.3	5'-CTA-ACT-GGT-ACC-ATA-ATT-TCA-CTA-AGG-GAG-G-3'	SEQ ID 4
Seq2B	5'-GGG-TCA-TAA-TAC-ACT-CCA-TG-3'	SEQ ID 18
Seq4B	5'-GGA-ATA-TTG-CTG-GTG-ATC-C-3'	SEQ ID 19
Seq6B	5'-CAT-TGT-TTA-ACT-TTT-GGG-CC-3'	SEQ ID 20
Seq7B	5'-GAT-AAA-ACC-TCC-AAT-TCC-3'	SEQ ID 21
Seq4A	5'-GTA-CAG-AAA-TGG-AAA-AGG-3'	SEQ ID 22
Seq6A	5'-GGA-TGA-TTT-GTA-TGT-AGG-3'	SEQ ID 23
Seq5B	5'-GGA-TGT-GGT-ATT-CCT-AAT-TG-3'	SEQ ID 24

Table 2 : Replacement or secondary sequencing primers. Initial preferred sequencing primers can be replaced by a set of possible replacement primers. Suggestions are indicated in the table.

Initial sequencing primer	Preferences set of replacement sequencing primers
Seq1FOR	Seq1F & Seq2A
Seq2FOR	Seq3A & Seq5A
Seq3F	Seq5A & Seq7A
Seq1B	PCR2.3 & Seq 2B
Seq3B	Seq2B & Seq4B
Seq6R	Seq6B & Seq7B

Modes for carrying out the invention:

A. *Amplification of the HIV-1 Protease -Reverse transcriptase coding domain*

RNA was isolated from 100 µl of plasma according to the method described by Boom et al., *J. Clin. Microbiol.* 28 (3) 495-503 (1990), and reverse transcribed with the GeneAmp reverse transcriptase kit (Perkin Elmer) as described by the manufacturer using a HIV-1 specific downstream primer (OUT3, see Table 1). Two subsequent nested PCR were set up using specific outer primers (PRTO-5 and OUT3) and inner primers (PCR2.5 and PCR2.3), respectively (see Table 1). The outer primer reaction was done as described in WO97/27480. The inner amplification was performed in a 96 well plate as follows: 4 µl of the outer amplification product was diluted to a final volume of 50 µl using a 10X amplification mix consisting of 5 µl 10X PCR buffer containing 15 mM MgCl₂, 1 µl dNTP's (10 mM) 0.5 µl PCR2.5 (0.25 µg/ml), 0.5 µl PCR2.3 (0.25 µg/ml), 0.4 µl Expand High Fidelity (3.5 U/µl) and MQ water. Amplification was initiated after a short denaturation of the amplification product made using the outer primers (2 min at 94 °C). 10 amplification cycles were started consisting of a 15 sec denaturation step at 94 °C, a 30 sec annealing step at 60 °C and a 2 min polymerase step at 72 °C, respectively. This amplification was immediately followed by 25 cycles consisting of a 15 sec denaturation step at 94 °C, a 30 sec annealing step at 60 °C and a x min polymerase step at 72 °C, respectively; where x started at 2 min and 5 sec and increased each cycle with 5 sec. Amplification was finalized by an additional polymerase step (7 min at 72 °C). Subsequently, the reaction was held at 4 °C till further analyzed or stored at -20 °C (for short periods) or -70 °C (for longer periods). In order to analyze the amplification products, a DNA agarose gel was run and amplification products were visualized using UV-detection. Obtained PCR products were purified using the QIAquick 96-well plate system as described by the manufacturer (Qiagen).

B. *Sequencing of pol coding region*

The coding domain of the *pol* gene present on the amplified fragments was analyzed via sequencing using standard sequencing techniques. Preferentially, one started initial with

a set of 6 primers (Seq1FOR, Seq2FOR, Seq3F, Seq1B, Seq3B and Seq6R) covering the coding domain of the HIV-protease and reverse transcriptase protein. Sequences and location onto the coding region are shown in Table 1 and Figure 1, respectively. The sequencing was started by first distributing 4 µl of the primer stocks (4.0 µM) over a 96 well plate where each stock is pipetted down the column. In a second step, master mixes were made consisting of 14 µl MQ, 17.5 µl dilution buffer, 7µl sample (PCR fragment) and 14 µl Big Dye Terminator Mix. A fraction (7.5 µl) of each master mix, containing a specific PCR fragment, was transferred to a specific place into the 96 well plate so that each sample fraction was mixed with a different PCR primer set. Samples were pipetted across the rows. Samples were placed in a thermal cycler and sequencing cycles started. The sequencing reaction consisted of 25 repetitive cycles of 10 sec at 96 °C, 5 sec at 50 °C and 4 min at 60 °C, respectively. Finally, sequence reactions were held at 4 °C till further analysis or stored as previously described. The sequencing reactions were precipitated using a standard ethanol precipitation procedure, resuspended in 2 µl formamide and heated for 2 minutes at 92 °C in the thermal cycler. Samples were cooled on ice until ready to load. 1 µl of each reaction was loaded on a 4.25% vertical acrylamide gel in a 377 sequencer system and gel was run until separation of the fragments was complete.

C. *Sequence analysis of pol coding region*

Sample sequences were imported as a specific project into the sequence manager of Sequencher (Genecodes) and compared to the wild type HXB2 Pro/RT reference sequence. Sequences were assembled automatically and set at 85% minimum match. Secondary peaks were searched and the minimum was set at 60%. Any sequence that hung over the 5' end or the 3' end of the reference was deleted. When a region of overlap between sequences from the same strand was reached, the poorest quality of sequence was deleted leaving an overlap of 5-10 bases. Ambiguous base calls are considered poor matches to exact base calls. The sequence assembly was saved within a contig that can be edited.

Obtained sequences were edited so that base calls could be interpreted easily. Ambiguous sequences were retrieved and checked for possible errors or points of heterogeneity. When the point of ambiguity appeared correct (both strands of sequence agree but is different from the reference sequence) it was interpreted to be a variant. The reference sequence was used as an aid for building a contig and a guide to overall size and for trimming, but was not used for deciding base calls. A change was only made when both strands agreed. All gaps were deleted or filled, unless they occur in contiguous groups of a multiple of 3 (I.E. insertion or deletion of complete codons) based on data from both sequence strands. Once the editing was complete, the new contig sequence was saved as a consensus sequence and used for further analysis.

Detailed sequence editing was performed following certain rules: A) ABI primer blobs are trimmed at 5' ends where 1 consecutive base remain off the scale; sequence is trimmed not more than 25% until the first 25 bases contain less than 1 ambiguity; at least first 10 bases from the 5' end are removed, B) 3' ends are trimmed starting 300 bases after the 5' trim; the first 25 bases containing more than 2 ambiguities are removed; trim from 3' end until the last 25 bases contain less than 1 ambiguity. The maximum length of the obtained sequence fragment after trimming is 550 bases.

Sequences that failed to align were removed from the assembly and replaced by data retrieved from new sequence analyses. When further failures occurred, PCR reactions were repeated. Chromatograms were visualized using the IBM software system.

D. Detection of clonal clinical samples- analysis of limit of detection for heterozygous base calls

A clonal clinical sample was mixed with wild type HXB2 at known ratio's to determine limits of detection of the system. The limit of detection was found to be around 1000 RNA copies/ml from plasma; mixed populations of mutations could be detected when present down to 20%.

It will be apparent to those skilled in the art that various modifications and variations can be made in the compositions and methods of the present invention without

-16-

Sub-B3